Health Canada Santé Canada

Healthy Environments and Consumer Safety

Development and Use of Predictive Hazard Modeling in the Categorization and Screening of Existing Substances at Health Canada

International Science Forum on Computational Toxicology

May 21st – 23rd, 2007

Research Triangle Park, NC 27711

Presented by: M.E. (Bette) Meek
Safe Environments Programme
Health Canada

Canada

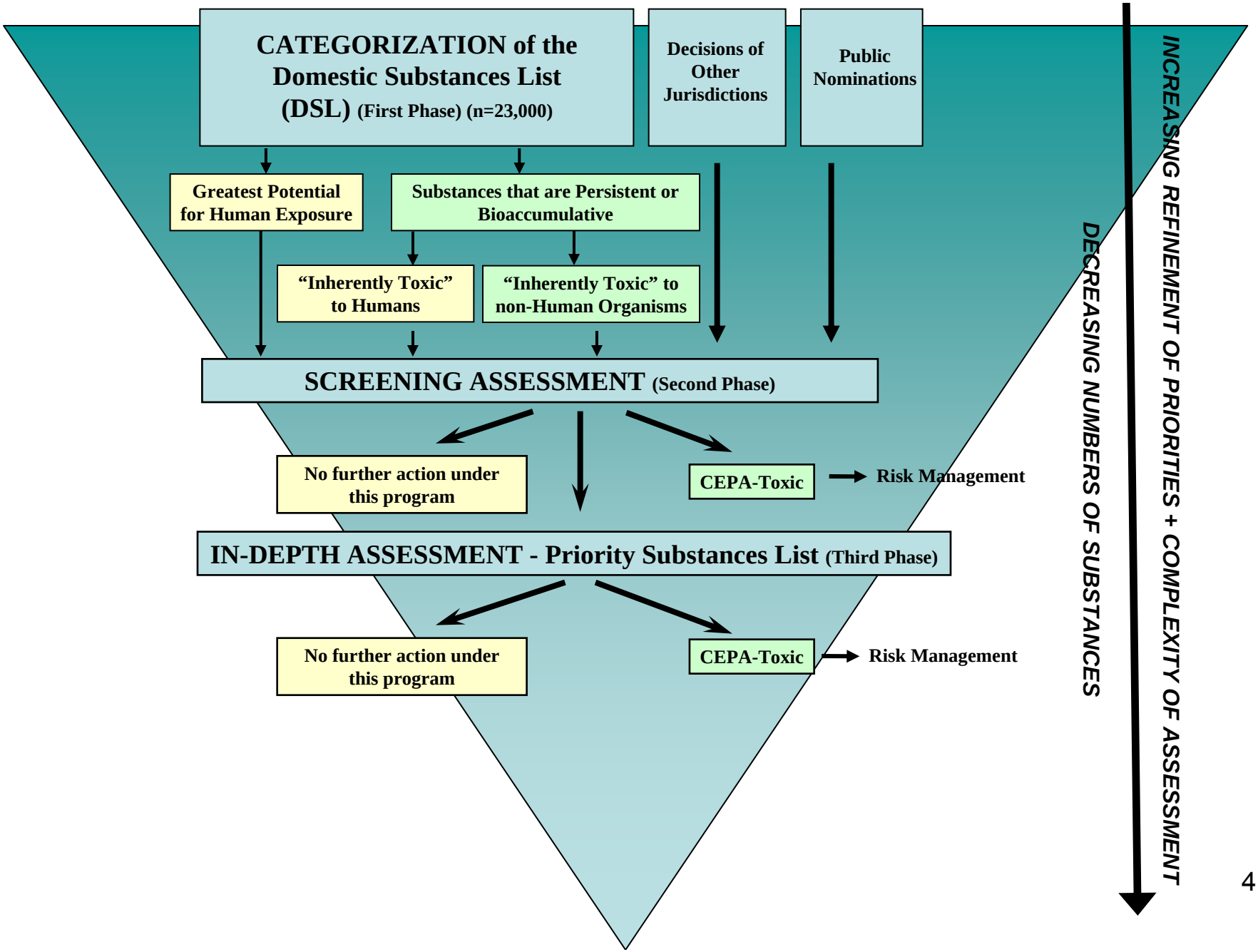
Outline

- Background – How predictive hazard tools contribute to the expanded mandate
- Weight of Evidence Approach for Cancer/Genotoxicity
  - Data/(Q)SAR/Analogues
- Robust Summaries for (Q)SAR Model Results
- Additional Development
- What we've learned
- Recommendations

Assessment of Existing Substances under the Canadian Environmental Protection Act (CEPA) – the Mandate for Human Health

- Address both exposure and effect to set priorities for risk management
  - consumer and environmental exposure
  - all media
- Publicly accountable – transparent process, documented outcome
- Under CEPA '88, assessments focussed on limited numbers of Priority Substances (n= 44 on PSL 1 and 25 on PSL 2)
- CEPA '99 extended our mandate to all Existing Substances in Canada (n=23,000)

CEPA 1999 Existing Substances Program



Simple and Complex Priority Setting Tools

EXPOSURE

Simple Exposure Tool (SimET) - Relative ranking of all DSL substances based on submitters (S),quantity (Q) and expert ranked use (ERU)

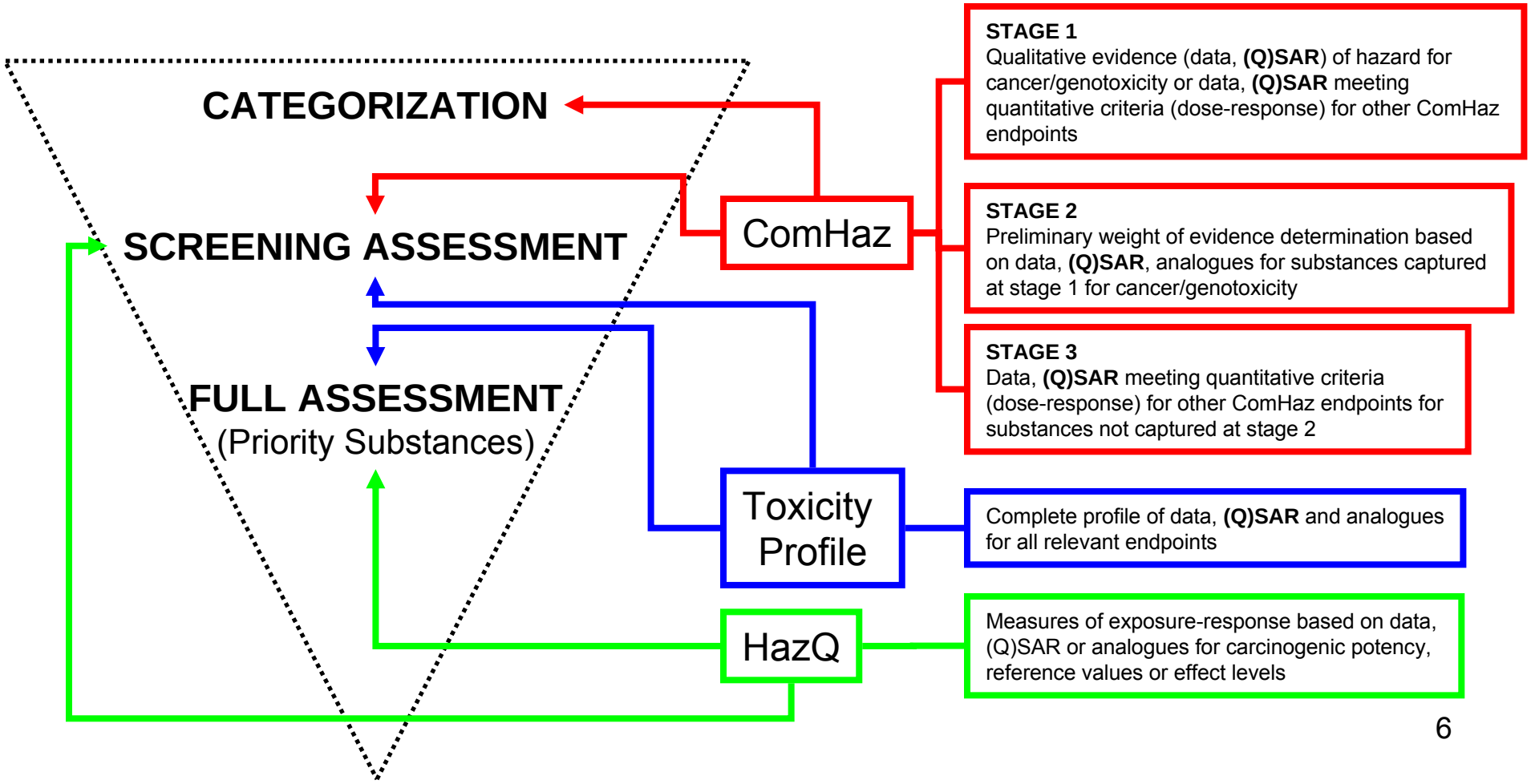
Complex Exposure Tool (ComET) - Quantitative plausible maximum age-specific estimates of environmental and consumer exposure for individuals based on use scenario (sentinel products), phys/chem properties & bioavailability

HAZARD

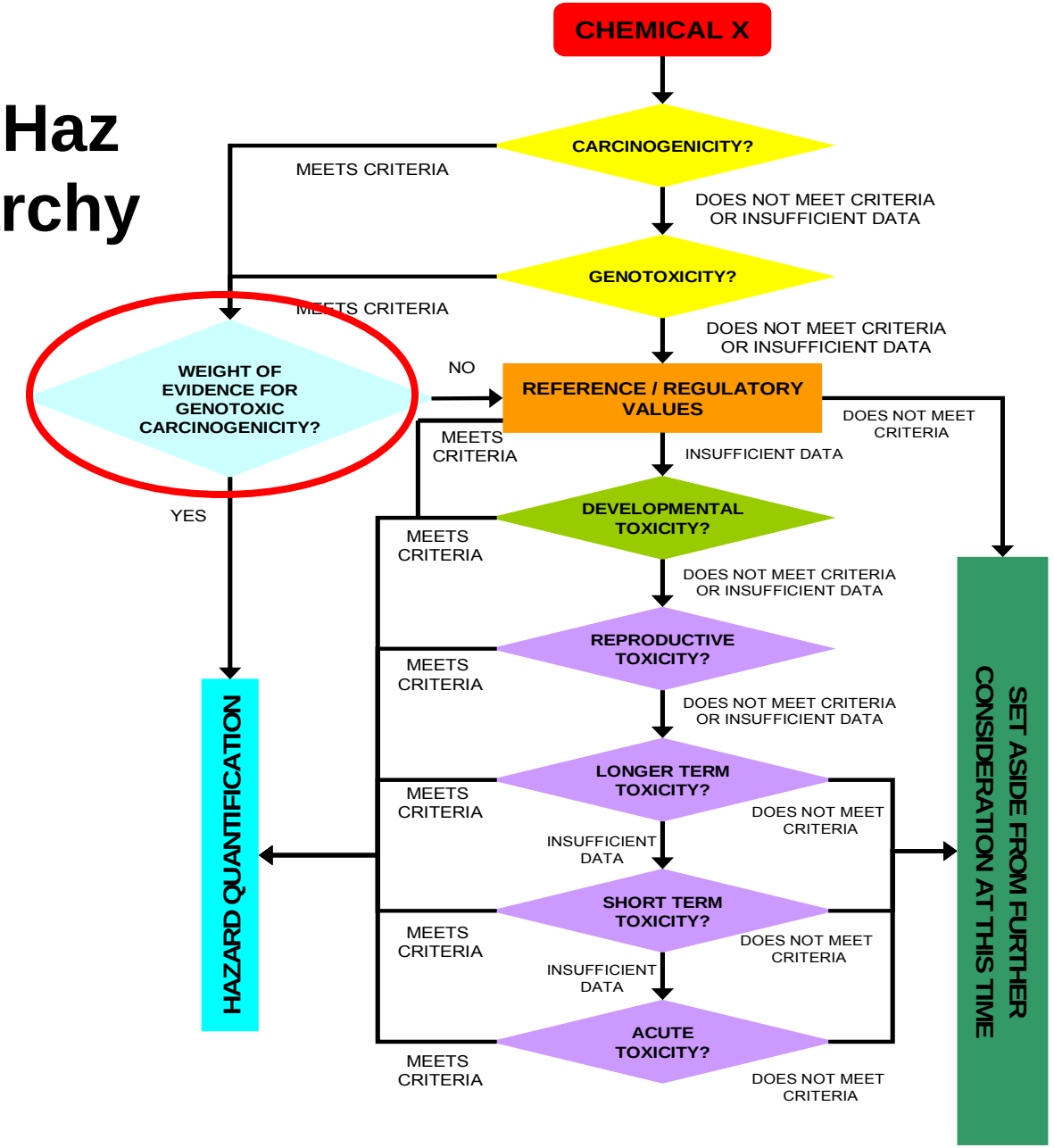
Simple Hazard Tool (SimHaz) - Identification of high or low hazard compounds by various agencies based on weight of evidence and expert opinion/consensus

Complex Hazard Tool (ComHaz) - Hierarchical approach for multiple endpoints & data sources (e.g., (Q)SAR) including preliminary weight of evidence framework

How (Q)SAR Contributes to Prioritization and Assessment of Existing Substances



ComHaz Hierarchy



Preliminary Weight of Evidence – Data/(Q)SAR/Analogues - Cancer/Genotoxicity

- Why cancer/genotoxicity?
  - High capture rate on first pass
    - Limited qualitative evidence sufficient
  - Confidence for predictive tools greatest for these endpoints
    - Larger more diverse training sets
    - Potential to consider endpoints in combination
    - Relevance to specific modes of action
      - E.g. frameshift mutation
      - Base pair substitution

(Q)SAR Models/Criteria for Selection

- Capability to generate predictions for a wide range of diverse chemical structures
  - Nature of the Domestic Substances List
- Capability to generate predictions for the endpoints of interest
- Ease of use and interpretation of results
- Transparency of basis for predictions

  e.g., information on critical chemical descriptors, coverage, prediction space, predictions for similar chemicals
- Computer requirements
- Availability and level of technical support

Preliminary Weight of Evidence Framework

(Q)SAR Models

CARCINOGENICITY (Q)SAR MODELS

TOPKAT* (version 6.2)

- **NTP Carcinogenicity**
  - Male Rat (195 chemicals) (v 3.2)
  - Female Rat (162 chemicals) (v 3.2)
  - Male Mouse (208 chemicals) (v 3.2)
  - Female Mouse (229 chemicals) (v 3.2)

CASETOX** (version 1.56) (Multicase PC Version)

- **NTP Rodent Carcinogenicity** (A07) (v 1.5) (313 chemicals)
- **FDA-CDER Rodent Carcinogenicity** (proprietary)
  - Male Rat (AG3) (v 1.7) (1301 chemicals)
  - Female Rat (AG4) (v 1.7) (1293 chemicals)
  - Male Mouse (AG1) (v 1.7) (1182 chemicals)
  - Female Mouse (AG2) (v 1.7) (1189 chemicals)

DEREK FOR WINDOWS*** (version 8.0.1)

- **49 Structural Alerts for Carcinogenicity**

GENOTOXICITY (Q)SAR MODELS)

TOPKAT* (version 6.2)

- **Ames Mutagenicity** (v 3.1) (1835 chemicals)

CASETOX** (version 1.56) (Multicase PC Version)

- **Salmonella Mutagencity (A2H) (v 1.7) (5864 chemicals)**
- **Somatic Mutations in Drosophila (A2D) (v 1.7) (289 chemicals)**
- **Mutations in Mouse Lymphoma Cells In Vitro (A2F) (v 1.5) (270 chemicals)**
- **Chromosomal Aberrations in CHO cells In Vitro (A61) (v 1.7) (805 chemicals)**
- **Induction of Micronuclei in Mouse Bone Marrow In Vivo (A62) (v 1.5) (238 chemicals)**
- **Unscheduled DNA Synthesis in Rat Hepatocytes In Vitro (A64) (v 1.5) (299 chemicals)**

DEREK FOR WINDOWS*** (version 8.0.1)

- **97 Structural Alerts for Genotoxicity**

*** Accelrys. Burlington, MA USA**
**** Multicase Incorporated. Beachwood, OH USA**
***** LHASA Ltd. Leeds, UK**

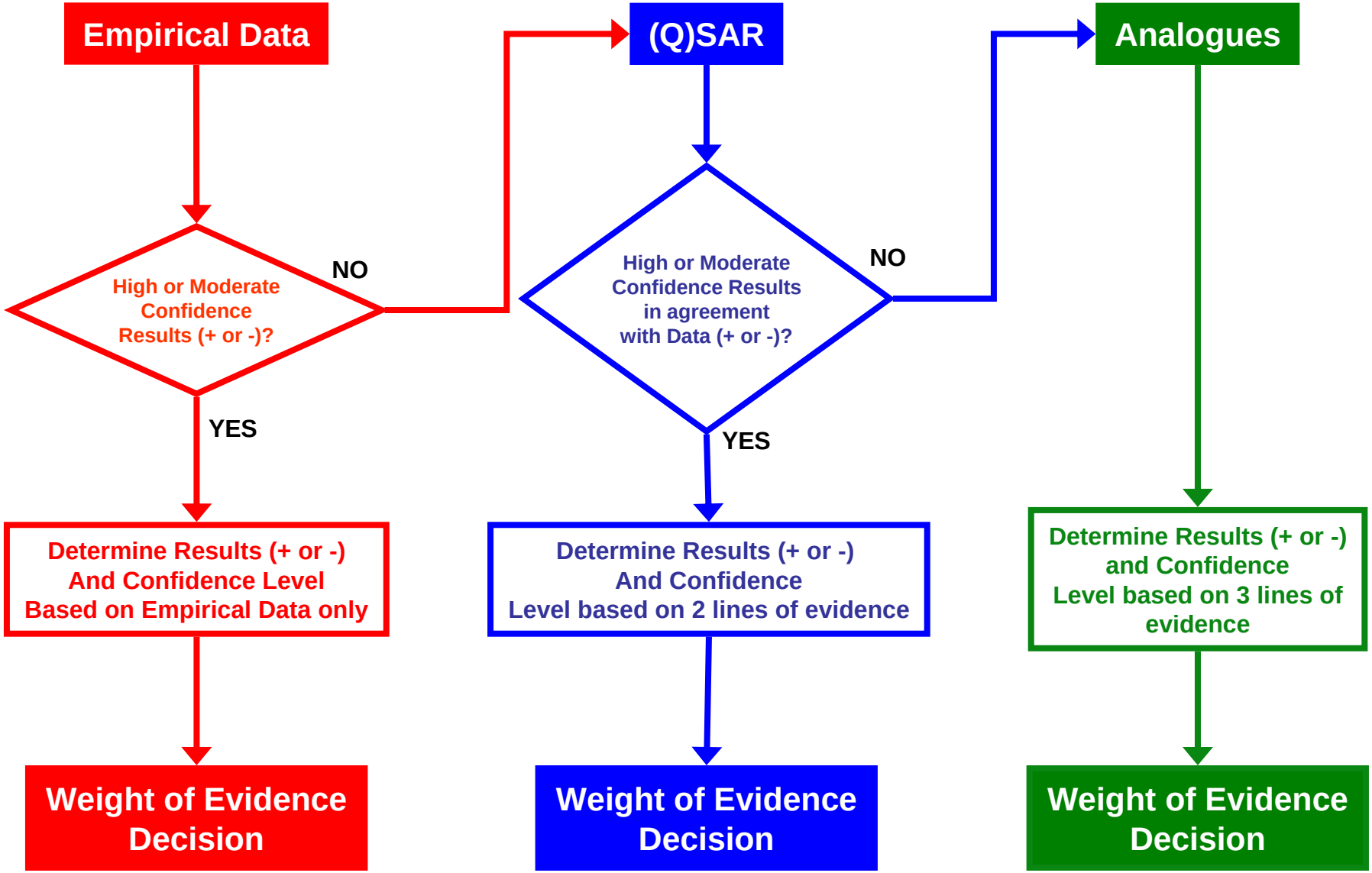
Principles – Development of the Preliminary Weight of Evidence Framework

- Relevance of/experience with transparent analytical frameworks which contribute to understanding and consistency
  - E.g., mode of action frameworks
  - (Meek et al., 2003; Seed et al., 2005; Boobis et al., 2007; in press)
- Criteria for consideration of the weight of evidence of output from predictive tools are similar to those for data:
  - E.g., strength, consistency, specificity, biological plausibility
- Need for transparency of process
  - Peer input, consultation & review
    - Complexity tailored to nature of decision

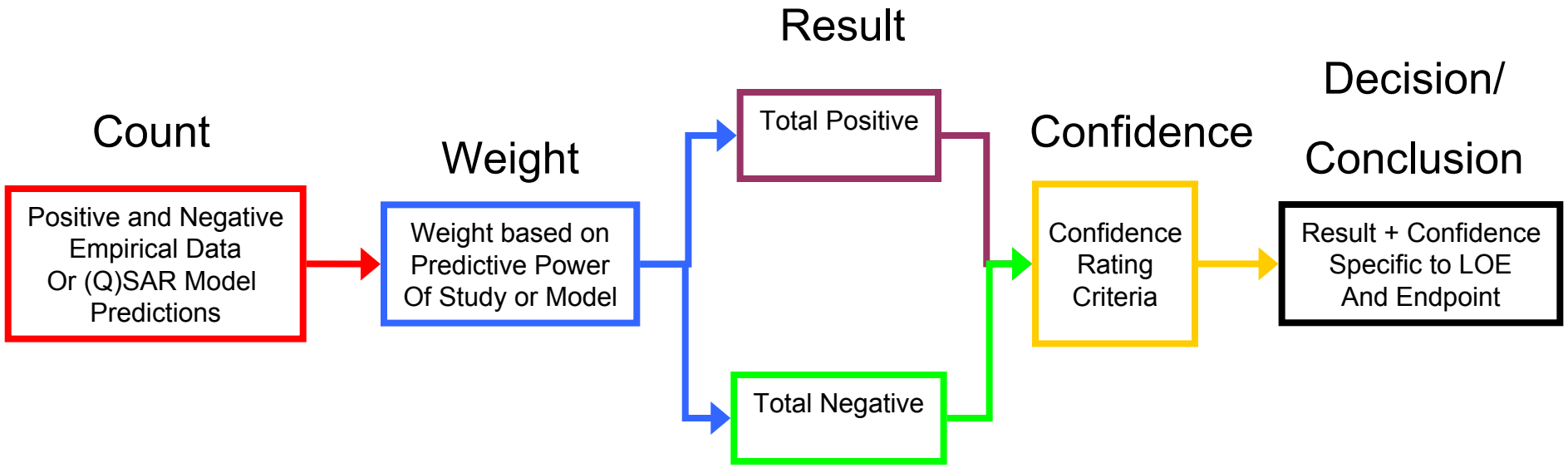
Preliminary Weight of Evidence Framework Proposal (Data/(Q)SAR/Analogues) - Process

- Draft approach developed/refined based on:
  - Previous external consultation with genetox experts
    - Weighting of predictive power of various bioassays
  - operational experience (hundreds of compounds)
  - Continuing consultation with internal genetox specialists
  - External peer consultation of modellers & endpoint specialists (mid 05)

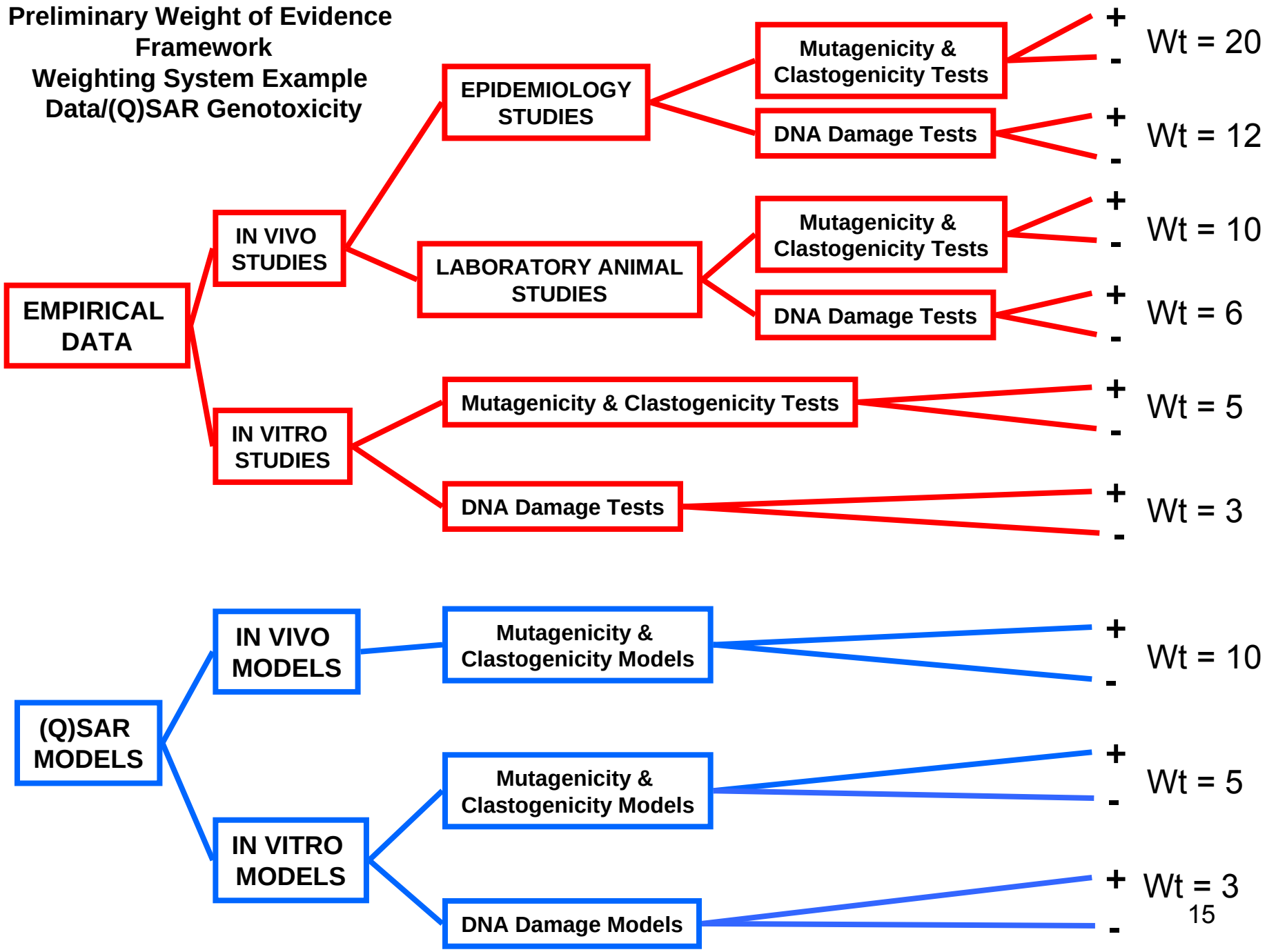
PRELIMINARY WEIGHT OF EVIDENCE FRAMEWORK



Basic Approach for each Line of Evidence (LOE)



Preliminary Weight of Evidence Framework
Weighting System Example
Data/(Q)SAR Genotoxicity



Analogue Component - Selection of Relevant Tools

Range of tools considered based on:

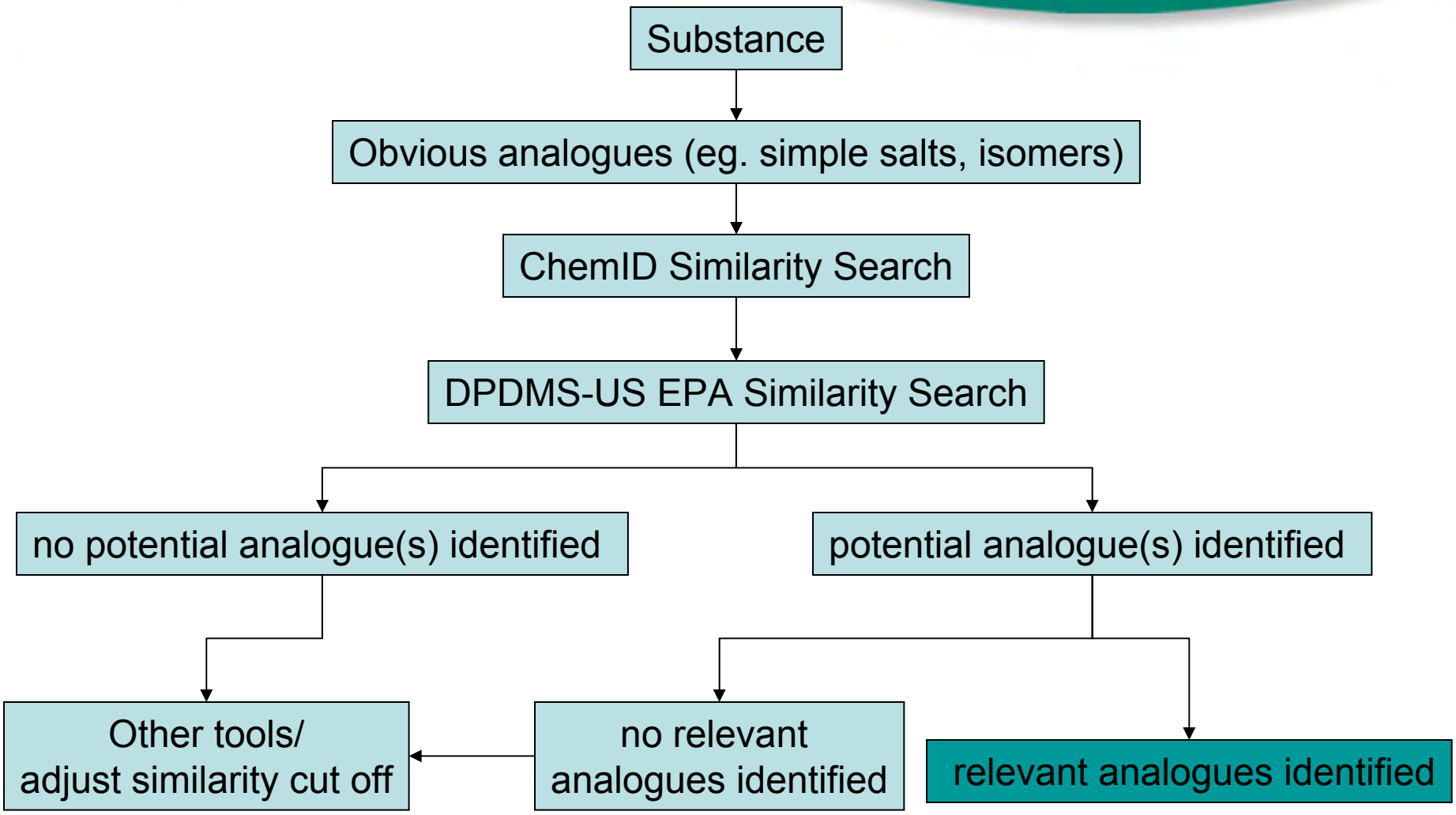
- Size of database + no. of structural features
- Ability to modify search criteria to limit no. of analogues identified
- Quantitative measures of tool peformance (e.g., no. of "good" analogues based on expert judgement/total no. of analogues for all tools)
- Ability to read-across toxicity data
- Ability to read-across physical-chemical data

Applied to a number of examples

Analogue Tools Applied

Tools	Size	Criteria
DPDMS	12 849 **(DSL)** 40,224 **(US EPA)** Structural features/fragments Accord 5.0 search engine + HC database	**% similarity, substructure** , equivalence + exact match
ChemID	247 407 960 fragments and weighting system	**% similarity, substructure** searches
Leadscope	25 000 fragments	**% similarity, substructure** searches
AIM	31 031 645 fragments, correction factors, ring index	**Qualitative similarity**
Topkat	165-1866 Descriptors molecular bulk, shape, symmetry and electrotopological state	**Qualitative similarity distance by endpoint**

Proposed Hierarchy



Robust Summaries of Modelled Results

- What?
  - Comparable to robust summary for data but including information relevant to interpretation of model output
  - Clear delineation of input and output of the model
  - Includes both chemical-specific and model-generic information
    - TOPKAT, Multicase, DEREK for Windows
- Why?
  - Critically important to transparency, increased understanding and capacity building
  - Provides information to assist in interpreting conflicting model predictions, difficult wt of evidence calls, etc.

Robust Summary of (Q)SAR Model Results

- Chemical structure, SMILES code
- Model identification (name, version, submodel)
- Size of database, distribution of pos/neg cmpds, performance data (model developer)
- Source of studies for database and criteria for inclusion
- Qualitative (+/- and/or probabilities) or quantitative (e.g., LD50) outputs
- Results of analyses of coverage, descriptor space, similar compounds
- Criteria for interpretation of predictions (e.g., model developer and/or Health Canada)
- Content consistent with OECD principles of consideration of a QSAR model for regulatory purposes

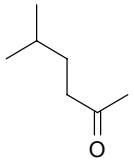
Robust Summaries of (Q)SAR Model Results – Example

Robust (Quantitative) Structure Activity Relationship ((Q)SAR) Model Summary

Evaluator: XXXX
Date: November 7, 2005

CHEMICAL INFORMATION

Chemical Name: 2-Hexanon, 5-methyl-
CAS No.: 110-12-3
SMILES Code: O=C(CCC(C)C)C
Structure :



COMPUTER PROGRAM AND MODEL

Program : TOPKAT Version 6.2 (2004)
Company : Accelrys Inc., San Diego, CA, USA
Program Type: TOPKAT is a statistically based QSAR system where predictions of the molecule are generated using descriptors that quantify molecular bulk, shape, symmetry and electropological state (E-state).

Model: NTP Carcinogenicity
Submodels: Male Rat (version 3.1)
Female Rat (version 3.1)
Male Mouse (version 3.1)
Female Mouse (version 3.1)

SUMMARY OF PREDICTIONS

Male Rat: Inconclusive
Female Rat: Inconclusive
Male Mouse: Negative
Female Mouse: Negative

BACKGROUND INFORMATION ON MODEL

Criteria for Inclusion of Chemicals in Database (Accelrys, 2004a)

- Carcinogenicity studies conducted in F344 rats or B6C3F1 mice by the National Cancer Institute (NCI) and National Toxicology Program (NTP).
- Oral exposure via food or drinking water, generally for 2 years.
- The NCI/NTP examines data through an expert committee process. Each chemical experiment is conducted in both sexes of F344 rats or B6C3F1 mice via oral administration. Each strain/sex conclusion is assigned one of the following categories: Clear Evidence, Some Evidence, Equivocal Evidence, or No Evidence of Carcinogenic Activity.
- "Positive" chemicals included in the model database were those with NTP classifications of "Clear Evidence" or "Some Evidence" of carcinogenicity with a purity of at least 95% and impurities known.
- "Negative" chemicals included in the model database were those with NTP classification of "No Evidence" of carcinogenicity; negative chemicals tested in rats exposed for at least 100 weeks and mice exposed for at least 70 weeks.

Size of Database and Distribution of Chemicals (Accelrys, 2004b)

Submodel	Chemicals used in derivation of model		Total used in model	Chemicals in database, but not used in derivation of model		Total not used in model	Total Chemicals
	Positive	Negative		Positive	Negative		
Male Rat	87	108	195	2	4	6	201
Female Rat	63	99	162	1	1	2	164
Male Mouse	66	142	208	0	1	1	209
Female Mouse	74	155	229	6	2	8	237

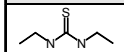
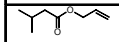
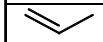
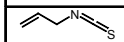
Internal Cross Validation (leave one out) Data for TOPKAT NTP Carcinogenicity Database (Accelrys, 2004a)

Submodel	Number of Chemicals	Specificity (%)	Sensitivity (%)	Indeterminate [Inconclusive] (%)
Male Rat	202	82	82	11
Female Rat	165	93	91	1
Male Mouse	210	94	90	1
Female Mouse	238	87	88	5

Example – continued …

Submodel: Male Mouse
Compound in database: No
Predicted Probability: 0.000
Univariate Analysis: Pass
Multivariate Analysis: Pass
Similarity Analysis: Pass; 6 similar compounds identified; all well predicted by model

Summary of Similarity Analysis for Male Mouse model

Structure	Chemical Name	CAS No	Similarity Distance	Used in model develop-ment (Yes/No)	Predicted Result	Actual Result	Reference
	N,N'-diethylthiourea	105-55-5	0.125	Yes	Neg	Neg	NTP TR 149
	Methyl methacrylate	80-62-6	0.150	Yes	Neg	Neg	NTP TR 314
	Trimethylphosphate	512-56-1	0.154	Yes	Neg	Neg	NTP TR 81
	Allylisovalerate	2835-39-4	0.185	Yes	Neg	Neg	NTP TR 253
	Propylene	115-07-1	0.191	Yes	Neg	Neg	NTP TR 272
	Allyl isothiocyanate	57-06-7	0.197	Yes	Neg	Neg	NTP TR 234

Summary of Contribution of Descriptors to the Male Mouse Model

Descriptor Fragment	Descriptor	Contribution
na	Symmetry Index #7	9.463
	[Aliphatic C]	4.504
	[-CH2-]	1.286
	[*C(=*)(*)]	-2.271
	[Aliphatic O]	-3.027
	[-CH3]	-4.732
na	Constant Term	-22.921

CRITERIA FOR INTERPRETATION OF PREDICTIONS

Criteria for conclusive predictions

Validation Analysis	Criteria	Description
Univariate Analysis	Pass	Automated; Search for coverage of structural fragments (1-2 atom sections of the queried structure) in the Training Set of the model.
Multivariate Analysis	Pass	Automated; Check on whether query structure falls within or near the multidimensional optimum prediction space (OPS) formed by model descriptors.
Similarity Analysis	Pass	User driven; Comparison of compounds in database to query structure based on values for structural descriptors; Similarity based on 0 – 1 scale (0 = 100% similar, 1 = totally dissimilar, ≤ 0.2 = minimum acceptable similarity); Compare predicted and actual empirical data results for similar compounds; Predicted and actual results must agree for ≥ 50% of similar compounds. Similarity Searching conducted using all descriptors in the model, therefore, the complete database. Alternatively, the user can search for similarity on a specific fragment.

Criteria for interpreting predictions

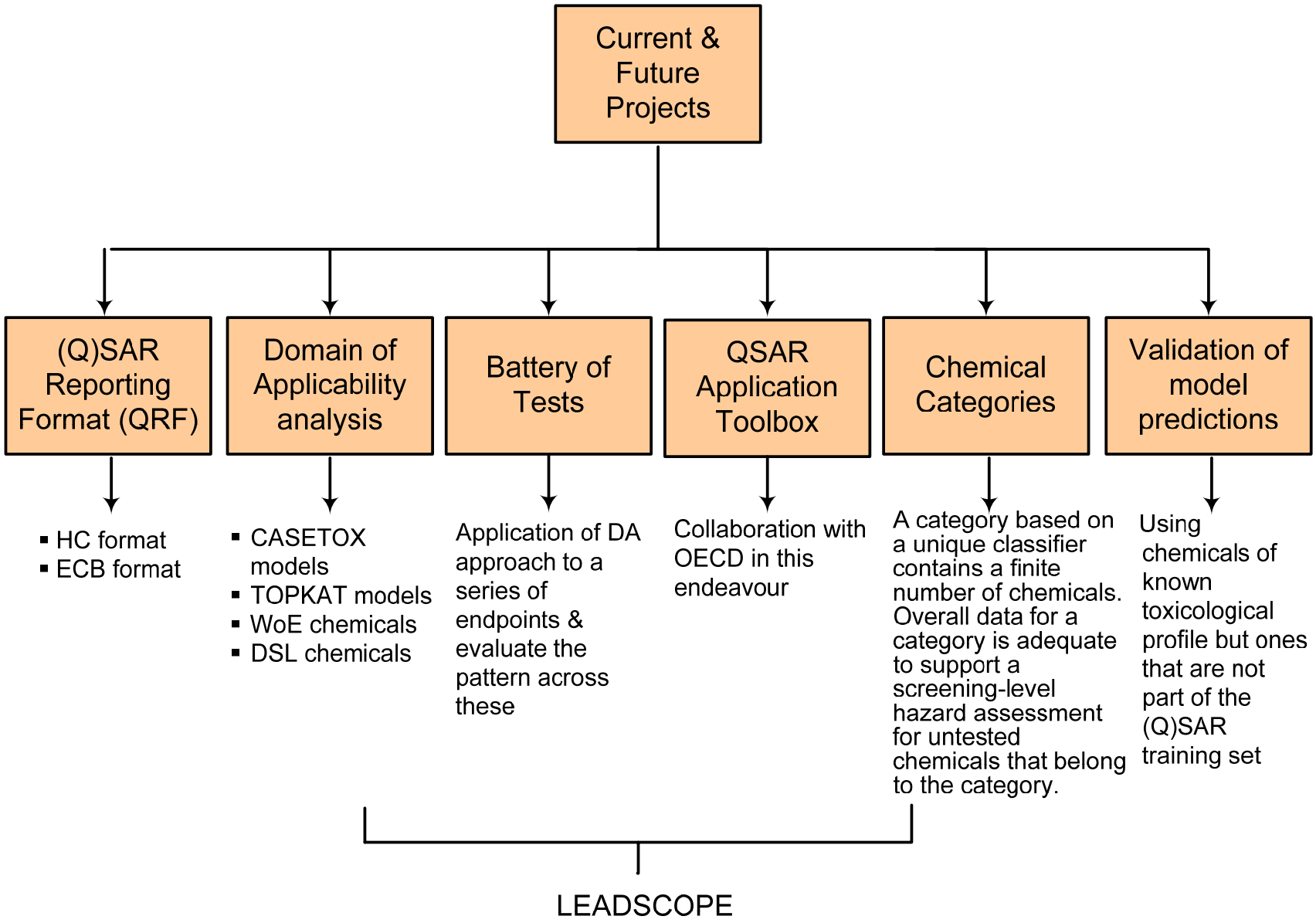
Predicted Probability	Conclusion	Equivalent NTP/NCI Classification
< 0.300	Negative	No Evidence
0.300 - 0.700	Indeterminate/Inconclusive	N/A
> 0.700	Positive	Clear Evidence, Some Evidence

ILSI Risk Science Institute Project on Improving the Predictive Base for Developmental Toxicity

- Sponsored by Health Canada
- Working Group includes model developers, developmental toxicologists (>20, currently)/biologists, risk assessors
- Objective: More systematic mining of data on developmental toxicity as a basis potentially for predictive tools

Objectives of the Database

- Compile essential data in a rigorous, transparent, objective and reproducible manner
- Compile findings on discrete endpoints, as reported
- Capture objective data, as opposed to high-level summary judgments
- Allow flexibility to examine alternative approaches to grouping data on specific endpoints
- Protoype available
  - Benz et al. poster at this meeting:
    - Towards Refined Use of Toxicity Data in Statistically Based SAR Models for Developmental Toxicity


Current & Future Projects
(Q)SAR Reporting Format (QRF)
Domain of Applicability analysis
Battery of Tests
QSAR Application Toolbox
Chemical Categories
Validation of model predictions
HC format
ECB format
CASETOX models
TOPKAT models
WoE chemicals
DSL chemicals
Application of DA approach to a series of endpoints & evaluate the pattern across these
Collaboration with OECD in this endeavour
A category based on a unique classifier contains a finite number of chemicals. Overall data for a category is adequate to support a screening-level hazard assessment for untested chemicals that belong to the category.
Using chemicals of known toxicological profile but ones that are not part of the (Q)SAR training set
LEADSCOPE

QSAR for DSL/Health – What We've Learned

- The value of (Q)SAR in contributing to weight of evidence decisions for human health
- The need for transparency in building credible weight of evidence approaches and building capacity and acceptance
  - Exploiting the greater comfort level with data and analogue approaches
- The need for transparent, publically available, thoughtfully developed models for broader application for human health endpoints
  - collaboration of risk assessors, endpoint specialists, modellers
  - Improved capture of relevant data
  - More mode of action relevance for human health

Recommendations: Increasing the Use of Computational Toxicology Models in Human Health Risk Assessment

- Communication/Consultation
  - early and continuing engagement of endpoint specialists/risk assessors in model development
  - Providing the basis to answer the right questions
- Transparency
  - Fully transparent, publically available
- Education and training
  - Earlier, more quantitative training of toxicologists
  - specialized courses for risk assessors/managers
- Applicability
  - Better characterization of chemical space through targetted testing strategies

More Information?

- **Health Canada Existing Substances Division Website** – http://www.hc-sc.gc.ca/ewh-semt/contaminants/existsub/index_e.html
- Report of the Peer Consultation on the Weight of Evidence Approach for Cancer/Genotoxicity
  - **www.tera.org**
- Refined use of developmental toxicity data for statistically based SARs
  - **rsi.ilsi.org**
- **Health Canada Existing Substances Mailing List** – http://www.hc-sc.gc.ca/ewh-semt/contaminants/existsub/mail-avis_e.html